Capacity Bounds for State-Dependent Broadcast Channels

K. G. Nagananda, Chandra R. Murthy and Shalinee Kishore*

Abstract

In this paper, we derive information-theoretic performance limits for three classes of two-user state-dependent discrete memoryless broadcast channels, with noncausal side-information at the encoder. The first class of channels comprises a sender broadcasting two independent messages to two non-cooperating receivers; for channels of the second class, each receiver is given the message it need not decode; and the third class comprises channels where the sender is constrained to keep each message confidential from the unintended receiver. We derive inner bounds for all the three classes of channels. For the first and second class of channels, we discuss the rate penalty on the achievable region for having to deal with side-information. For channels of third class, we characterize the rate penalties for having to deal not only with side-information, but also to satisfy confidentiality constraints. We then derive outer bounds, where we present an explicit characterization of sum-rate bounds for the first and third class of channels. For channels of the second class, we show that our outer bounds are within a fixed gap away from the achievable rate region, where the gap is independent of the distribution characterizing this class of channels. The channel models presented in this paper are useful variants of the classical broadcast channel, and provide fundamental building blocks for cellular downlink communications with side-information, such as fading in the wireless medium, interference caused by neighboring nodes in the network, *etc.* at the encoder; two-way relay communications; and secure wireless broadcasting.

Keywords: State-dependent broadcast channels, side-information, rate regions, outer bounds.

1 Introduction

The information-theoretic study of broadcast channels (BC) was initiated first by Cover in [1]. In the classical setting, the BC comprises a sender who wishes to transmit k independent messages to

*K. G. Nagananda and Shalinee Kishore are with the Dept. of ECE at Lehigh University, Bethlehem, PA, U.S.A. E-mail: {kgn209,skishore}@lehigh.edu; Chandra R. Murthy is with the Dept. of ECE at the Indian Institute of Science, Bangalore, India. E-mail: cmurthy@ece.iisc.ernet.in. Corresponding author: K. G. Nagananda.

k noncooperative receivers. The largest known inner bound on the capacity region when $k = 2$ was derived by Marton [2]. Recently, some ideas were discussed in [3], that is conjectured to lead to a larger inner bound. Capacity outer bounds were presented by Sato in [4] by utilizing the fact that the capacity region of BC depends on the marginal transition probabilities. Nair and El Gamal provided outer bounds for the two-user case [5], based on the results of the more capable BC [6]. Liang *et. al* generalized the outer bounds of [5] by deriving the *New-Jersey* outer bound. Some properties of the *New-Jersey* outer bound were exposed in [7], where it was shown to be equivalent to the computable UVW-bound with bounded cardinalities of the auxiliary random variables.

Several variants of this classical setting have also received considerable attention. One of the most prominent variants is the state-dependent BC with side-information, where the probability distribution characterizing the channel depends on a state process, and with the channel state made available as side-information at the transmitter, or at the receiver, or at both ends. Capacity inner bounds for the two-user BC with noncausal side-information at the transmitter were derived in [8], where Marton's achievability scheme was extended to state-dependent channels. In [9], inner and outer bounds were derived for the degraded BC with noncausal side-information at the transmitter; the capacity region was derived when side-information was obtained to the encoder in a causal manner. The capacity region for BC with receiver side-information was derived in [10], where a genie provides each receiver with the message it need not decode. To the best of the authors' knowledge, outer bounds for the two-user BC with noncausal side-information at the encoder have not appeared in the literature.

Yet another issue in wireless communications, owing to the broadcast nature of the wireless medium, is related to information security. That is, the broadcast nature of wireless networks facilitates malicious or unauthorized access to confidential data, denial of service attacks, corruption of sensitive data, *etc.* An information-theoretic approach to address problems related to security has gained rapid momentum, and is commonly referred to as information-theoretic confidentiality or wireless physical-layer security [11]. An information-theoretic approach to secure broadcasting was inspired by the pioneering work of Csiszár and Körner [12], who derived capacity bounds for the two-user BC, when the sender transmits a private message to receiver 1 and a common message to both receivers, while keeping the private message confidential from receiver 2. Secure broadcasting with a single transmitter and multiple receivers in the presence of an external eavesdropper was considered in [13], where the secrecy capacity region was obtained for several special classes of channels. In [14], capacity bounds were derived for BC where a sender broadcasts two independent messages to two receivers, while keeping each message confidential from the unintended receiver. Capacity results and bounds for Gaussian BC with confidential messages were reported in [15] - [17]. The reader is referred to [18] for a comprehensive review of physical-layer security in BC. However, to the best of the authors' knowledge, the joint problem of side-information and confidentiality on the BC has not been addressed in the literature.

1.1 Main contributions

In this paper, we aim to provide useful insights into the effect of noncausal side-information at the encoder on (1) the classical two-user BC; (2) the BC with genie- aided receiver side-information; and (3) the BC with confidentiality constraints on the messages. Towards this end, we define three different classes of two-user discrete memoryless BC with noncausal side-information at the encoder. Of particular interest is the Class III channels (described below), which provides a fundamental building block to jointly address side-information and confidentiality in BC.

1. Class I: A sender broadcasts two independent messages to two non-cooperating receivers (see Fig. 1(a)). We derive an inner bound for this class of channels and characterize the rate penalty for dealing with noncausal side-information at the encoder. We are mainly concerned with outer bounds for this class of channels, where we present an explicit single-letter characterization of the sum-rate bound, along with bounds on single-user rates. An example for Class I channels is a base-station transmitting to two mobile receivers, with the base-station having prior knowledge of interference from a transmitter located in its vicinity, *e.g.*, through a backhaul network.

2. Class II: A sender broadcasts two independent messages to two receivers, with each receiver having *a priori* knowledge of the message it need not decode (see Fig. 1(b)). An example of this scenario is full-duplex communications between two nodes, aided by a relay. The relay node broadcasts the messages to the terminals, with each terminal knowing its own message. We devise an achievability scheme to derive an inner bound for this class of channels and show that the achievable rate for each user is in fact the maximum rate achievable for a single-user channel with states known *a priori* at the encoder. We also derive an outer bound which is within a fixed gap away from the achievable region, where the gap is independent of the

distribution characterizing this class of channels.

3. Class III: A sender broadcasts two independent messages to two receivers, such that each message is kept confidential from the unintended receiver (see Fig. 1(c)). To the best of the authors' knowledge, this is the first instance of a study of *simultaneous* impact of side-information and confidentiality constraints on BC. An inner bound for this class of channels is derived employing stochastic encoders to satisfy confidentiality constraints; we characterize the rate penalties for having to deal not only with side-information, but also to satisfy confidentiality constraints. One of the outer bounds is derived by employing a genie, which gives one of the receivers the message it need not decode, while the other receiver computes the equivocation rate treating this message as side-information. We also derive another outer bound, with an explicit characterization of the sum-rate bounds. As an example for this class of channels, we can extend the example considered for Class I channels, with the additional constraint of keeping each message confidential from the unintended receiver.

The remainder of the paper is organized as follows. In Section 2, we introduce the notation used and provide a mathematical model for the discrete memoryless version of the channels considered in this paper. In Section 3, we summarize the main results of this paper by describing inner and outer bounds for all the channel models, and provide related discussion. The proofs of the achievability theorems can be found in Section 4, while the proofs of the outer bounds are provided in Section 5. Finally, we conclude the paper in Section 6. The encoder error analysis is relegated to Appendix A.

2 System model and notation

The channels belonging to Class I, Class II and Class III are denoted C_1, C_2 and C_3, respectively. Calligraphic letters are used to denote finite sets, with a probability function defined on them. N is the number of channel uses, and $\mathrm{n} = 1, \ldots, \mathrm{N}$ denotes the channel index. Uppercase letters denote random variables (RV), while boldface uppercase letters denote a sequence of RVs. The following notation for a sequence of RVs is useful: $\mathbf{Y}_1^{\mathrm{N}} \triangleq (Y_{1,1}, \ldots, Y_{1,N})$; $\mathbf{Y}_1^{\mathrm{n}-1} \triangleq (Y_{1,1}, \ldots, Y_{1,\mathrm{n}-1})$; and $\mathbf{Y}_{1,\mathrm{n}+1}^{\mathrm{N}} \triangleq (Y_{1,\mathrm{n}+1}, \ldots, Y_{1,\mathrm{N}})$. Lowercase letters are used to denote particular realizations of RVs, and boldface lowercase letters denote vectors. The sender is denoted S and the receivers are denoted D_t, where $t = 1, 2$ is the receiver index. Discrete RV $X \in \mathcal{X}$ and $Y_t \in \mathcal{Y}_t$ denote the channel input and outputs, respectively. The encoder of S is supplied with side-information $\mathbf{W} \in \mathcal{W}^{\mathrm{N}}$, in a noncausal manner. The channel is assumed to be memoryless and is characterized by the conditional distribution $p(\mathbf{Y}_1, \mathbf{Y}_2|\mathbf{X}, \mathbf{W}) = \prod_{\mathrm{n}=1}^{\mathrm{N}} p(Y_{1,\mathrm{n}}, Y_{2,\mathrm{n}}|X_{\mathrm{n}}, W_{\mathrm{n}})$. For sake of brevity, in the remainder of this paper, we use $p(x)$ to denote $p(X = x)$. Unless otherwise stated, $p(\mathbf{x}) = \prod_{\mathrm{n}=1}^{\mathrm{N}} p(x_{\mathrm{n}})$.

To transmit its messages, S generates two RVs $M_t \in \mathcal{M}_t$, where $\mathcal{M}_t = \{1, \ldots, 2^{\mathrm{N}R_t}\}$ denotes a set of message indices. Without loss of generality, $2^{\mathrm{N}R_t}$ is assumed to be an integer, with R_t being the transmission rate intended to D_t. M_t denotes the message S intends to transmit to D_t, and is assumed to be independently generated and uniformly distributed over the finite set $\mathcal{M}_t$. Integer $m_t \in \mathcal{M}_t$ is a particular realization of M_t and denotes the message-index.

Given the conditional distribution characterizing the channel, a $((2^{\mathrm{N}R_1}, 2^{\mathrm{N}R_2}), \mathrm{N}, P_e^{(\mathrm{N})})$ code for the channels C_1 and C_2 comprises N encoding functions f, such that $\mathbf{X} = \mathbf{f}(m_1, m_2, \mathbf{W})$; for the channel C_3, it comprises a stochastic encoder, which is defined by the matrix of conditional probabilities $\phi(\mathbf{X}|m_1, m_2, \mathbf{W})$, such that $\sum_{\mathbf{X}} \phi(\mathbf{X}|m_1, m_2, \mathbf{W}) = 1$. Here, $\phi(\mathbf{X}|m_1, m_2, \mathbf{W})$ denotes the probability that a pair of message-indices (m_1, m_2) is encoded as $\mathbf{X} \in \mathcal{X}^{\mathrm{N}}$ to be transmitted by S, in the presence of noncausal side-information $\mathbf{W}$. For all channel models, there are two decoders $g_t : \mathcal{Y}_t^{\mathrm{N}} \to \mathcal{M}_t$.

The average probability of decoding error for the code, averaged over all codes, is $P_e^{(\mathrm{N})} = \max\{P_{e,1}^{(\mathrm{N})}, P_{e,2}^{(\mathrm{N})}\}$, where, $P_{e,t}^{(\mathrm{N})} = \sum_{m_1,m_2} \sum_{\mathbf{W} \in \mathcal{W}^{\mathrm{N}}} \frac{1}{2^{\mathrm{N}[R_1+R_2]}} \Pr\left[g_t(\mathcal{Y}_t^{\mathrm{N}}) \neq m_t | m_1, m_2, \mathbf{W}\right]$. A rate pair (R_1, R_2) is said to be achievable for the channel $\mathrm{C}_{\mathrm{c}}; \mathrm{c} = 1, 2, 3$, if there exists a sequence of $((2^{\mathrm{N}R_1}, 2^{\mathrm{N}R_2}), \mathrm{N}, P_e^{(\mathrm{N})})$ codes, such that $\forall \epsilon > 0$ and sufficiently small, $P_e^{(\mathrm{N})} \leq \epsilon$ as $\mathrm{N} \to \infty$. Furthermore, for the channel C_3, the following constraints [19] on the conditional entropy must be satisfied for (R_1, R_2) to be considered achievable:

$$\mathrm{N}R_1 - H(M_1|\mathbf{Y}_2) \leq \mathrm{N}\epsilon, \tag{1}$$

$$\mathrm{N}R_2 - H(M_2|\mathbf{Y}_1) \leq \mathrm{N}\epsilon. \tag{2}$$

The capacity region is defined as the closure of the set of all achievable rate pairs (R_1, R_2).

3 Main results

In this section, we state the achievability and converse theorems for all the channel models considered in this paper, and provide related discussion. Let $\mathcal{C}_{\mathrm{c}}$ denote the capacity region of the channel C_{c};

c $= 1, 2, 3$. We use the following auxiliary RVs defined on finite sets: $U \in \mathcal{U}$, $V_1 \in \mathcal{V}_1$ and $V_2 \in \mathcal{V}_2$.

3.1 Class I **channels**

For the channel C_1, we consider the set $\mathcal{P}_1$ of all joint probability distributions $p_1(.)$ that can be factored as $p(w)p(v_1, v_2|w)p(x|w, v_1, v_2)p(y_1, y_2|x)$. For a given $p_1(.) \in \mathcal{P}_1$, a lower bound on the capacity region for C_1 is described by the set $\mathcal{R}_{1,\mathrm{in}}(p_1)$, which is defined as the union over all distributions $p_1(.)$ of the convex hull of the set of all rate pairs (R_1, R_2) that simultaneously satisfy (3) - (5).

$$
\begin{aligned}
R_1 &\leq I(V_1; Y_1) - I(V_1; W), && (3)\\
R_2 &\leq I(V_2; Y_2) - I(V_2; W), && (4)\\
R_1 + R_2 &\leq I(V_1; Y_1) + I(V_2; Y_2) - I(V_1; V_2) - I(V_1, V_2; W), && (5)
\end{aligned}
$$

where V_1 and V_2 are constrained to satisfy the Markov chain $(V_1, V_2) \to (X, W) \to (Y_1, Y_2)$.

Theorem 3.1 *Let* $\mathcal{R}_{1,in} = \bigcup_{p_1(.) \in \mathcal{P}_1} \mathcal{R}_{1,in}(p_1)$. *Then,* $\mathcal{R}_{1,in} \subseteq \mathcal{C}_1$.

For proof, see Section 4.1.

For a given $p_1(.) \in \mathcal{P}_1$, an outer bound for C_1 is described by the set $\mathcal{R}_{1,\mathrm{out}}(p_1)$, which is defined as the union of all rate pairs (R_1, R_2) that simultaneously satisfy (6) - (7).

$$
\begin{aligned}
R_1 &\leq I(V_1; Y_1) - I(V_1; W), && (6)\\
R_2 &\leq I(V_2; Y_2) - I(V_2; W), , && (7)
\end{aligned}
$$

where $(V_1, V_2) \to (X, W) \to (Y_1, Y_2)$.

Theorem 3.2 *Let* $\mathcal{R}_{1,out} = \bigcup_{p_1(.) \in \mathcal{P}_1} \mathcal{R}_{1,out}(p_1)$. *Then,* $\mathcal{C}_1 \subseteq \mathcal{R}_{1,out}$.

The proof of Theorem 3.2 can be found in Section 5.1. However, this outer bound does not include a bound on the sum-rates. To explicitly bound the sum-rate, we provide the following alternative outer bound for the channel C_1. We consider the set $\mathcal{P}_1^*$ of all joint probability distributions $p_1^*(.)$ that can be factorized as follows: $p(w)p(u, v_1, v_2|w)p(x|w, u, v_1, v_2)p(y_1, y_2|x)$. For a given $p_1^*(.) \in \mathcal{P}_1^*$, an outer bound for C_1 is described by the set $\mathcal{R}_{1,\mathrm{out}}^*(p_1^*)$, which is defined as the union of all rate pairs

(R_1, R_2) that simultaneously satisfy (8) - (11).

$$R_1 \leq I(U, V_1; Y_1) - I(V_1; W|U), \tag{8}$$

$$R_2 \leq I(U, V_2; Y_2) - I(V_2; W|U), \tag{9}$$

$$R_1 + R_2 \leq I(U, V_1; Y_1) - I(V_1; W|U) + I(U, V_2; Y_2|V_1) - I(V_2; W|U, V_1), \tag{10}$$

$$R_1 + R_2 \leq I(U, V_2; Y_2) - I(V_2; W|U) + I(U, V_1; Y_1|V_2) - I(V_1; W|U, V_2), \tag{11}$$

where the following Markov chain is satisfied: $(U, V_1, V_2) \rightarrow (X, W) \rightarrow (Y_1, Y_2)$.

Theorem 3.3 *Let* $\mathcal{R}^*_{1,out} = \bigcup_{p^*_1(.)\in\mathcal{P}^*_1} \mathcal{R}^*_{1,out}(p^*_1)$. *Then,* $\mathcal{C}_1 \subseteq \mathcal{R}^*_{1,out}$.

Section 5.2 contains the proof of Theorem 3.3.

3.2 Class II **channels**

For the channel C_2, we consider the set $\mathcal{P}_2$ of all joint probability distributions $p_2(.)$ of the form $p(w)p(u|w)p(x|w,u)p(y_1,y_2|x)$. For a given $p_2(.) \in \mathcal{P}_2$, a lower bound on the capacity region for C_2 is described by the set $\mathcal{R}_{2,\mathrm{in}}(p_2)$, which is defined as the union over all distributions $p_2(.)$ of the convex-hull of the set of all rate pairs (R_1, R_2) that simultaneously satisfy (12) - (13).

$$R_1 \leq I(U; Y_1) - I(U; W), \tag{12}$$

$$R_2 \leq I(U; Y_2) - I(U; W), \tag{13}$$

where the Markov chain $U \rightarrow (X, W) \rightarrow (Y_1, Y_2)$ holds.

Theorem 3.4 *Let* $\mathcal{R}_{2,in} = \bigcup_{p_2(.)\in\mathcal{P}_2} \mathcal{R}_{2,in}(p_2)$. *Then,* $\mathcal{R}_{2,in} \subseteq \mathcal{C}_2$.

The proof of Theorem 3.4 is relegated to Section 4.2.

For a given $p_2(.) \in \mathcal{P}_2$, an outer bound for C_2 is described by the set $\mathcal{R}_{2,\mathrm{out}}(p_2)$, which is defined as the union of all rate pairs (R_1, R_2) that simultaneously satisfy (14) - (15).

$$R_1 \leq I(U; Y_1) - I(U; W) + H(U), \tag{14}$$

$$R_2 \leq I(U; Y_2) - I(U; W) + H(U), \tag{15}$$

with $U \rightarrow (X, W) \rightarrow (Y_1, Y_2)$.

Theorem 3.5 *Let* $\mathcal{R}_{2,out} = \bigcup_{p_2(.)\in\mathcal{P}_2} \mathcal{R}_{2,out}(p_2)$. *Then,* $\mathcal{C}_2 \subseteq \mathcal{R}_{2,out}$.

The proof of Theorem 3.5 can be found in Section 5.3.

3.3 Class III **channels**

For the channel C$_3$, we consider the set $\mathcal{P}_3$ of all joint probability distributions $p_3(.)$ that can be written as $p(w)p(u)p(v_1, v_2|w, u)p(x|w, v_1, v_2)p(y_1, y_2|x)$. For a given $p_3(.) \in \mathcal{P}_3$, an inner bound on the capacity region for C$_3$ is described by the set $\mathcal{R}_{3,\text{in}}(p_3)$, which is defined as the union over all distributions $p_3(.)$ of the convex-hull of the set of all rate pairs (R_1, R_2) that simultaneously satisfy (16) - (18).

$$\begin{aligned} R_1 &\leq I(V_1; Y_1|U) - \max[I(V_1; Y_2|U, V_2), I(V_1; W|U)], && (16)\\ R_2 &\leq I(V_2; Y_2|U) - \max[I(V_2; Y_1|U, V_1), I(V_2; W|U)], && (17)\\ R_1 + R_2 &\leq I(V_1; Y_1|U) + I(V_2; Y_2|U) - I(V_1; Y_2|U, V_2) - I(V_2; Y_1|U, V_1) \\ &\quad -I(V_1; V_2|U) - I(V_1, V_2; W|U), && (18) \end{aligned}$$

where the following Markov chain is satisfied: $U \to (V_1, V_2) \to (X, W) \to (Y_1, Y_2)$.

Theorem 3.6 *Let* $\mathcal{R}_{3,in} = \bigcup_{p_3(.)\in\mathcal{P}_3} \mathcal{R}_{3,in}(p_3)$. *Then,* $\mathcal{R}_{3,in} \subseteq \mathcal{C}_3$.

Section 4.3 contains the proof of Theorem 3.6.

For a given $p_3(.) \in \mathcal{P}_3$, an outer bound for C$_3$ is described by the set $\mathcal{R}_{3,\text{out}}(p_3)$, which is defined as the union of all rate pairs (R_1, R_2) that simultaneously satisfy (19) - (20).

$$\begin{aligned} R_1 &\leq \min[I_1, I_1^*], && (19)\\ R_2 &\leq \min[I_2, I_2^*],, && (20) \end{aligned}$$

where $I_1, \ldots, I_2^*$ are given by (21) - (24), respectively.

$$\begin{aligned} I_1 &\triangleq I(V_1; Y_1|U) - I(V_1; Y_2|U) + H(W|U, V_1), && (21)\\ I_2 &\triangleq I(V_2; Y_2|U) - I(V_2; Y_1|U) + H(W|U, V_2), && (22)\\ I_1^* &\triangleq I(V_1; Y_1|U, V_2) - I(V_1; Y_2|U, V_2) + H(W|U, V_1, V_2), && (23)\\ I_2^* &\triangleq I(V_2; Y_2|U, V_1) - I(V_2; Y_1|U, V_1) + H(W|U, V_1, V_2), && (24) \end{aligned}$$

where $U \to (V_1, V_2) \to (X, W) \to (Y_1, Y_2)$. The expressions (23) - (24) are obtained by letting a genie give D$_1$ message M_2, while D$_2$ computes the equivocation using M_2 as side-information.

Theorem 3.7 *Let* $\mathcal{R}_{3,out} = \bigcup_{p_3(.)\in\mathcal{P}_3} \mathcal{R}_{3,out}(p_3)$. *Then,* $\mathcal{C}_3 \subseteq \mathcal{R}_{3,out}$.

The proof of Theorem 3.7 can be found in Section 5.4. We also provide the following outer bound for the channel C_3, which explicitly characterizes the sum-rates. Consider the set $\mathcal{P}_3^*$ of all joint probability distributions $p_3^*(.)$ that can be factorized as follows: $p(w)p(u,v_1,v_2|w)p(x|w,u,v_1,v_2)$ $p(y_1,y_2|x)$. For a given $p_3^*(.) \in \mathcal{P}_3^*$, an outer bound for C_3 is described by the set $\mathcal{R}_{3,\mathrm{out}}^*(p_3^*)$, which is defined as the union of all rate pairs (R_1, R_2) that simultaneously satisfy (25) - (28).

$$\begin{aligned} R_1 &\leq I(U,V_1;Y_1) - I(V_1;W|U) - I(V_1;Y_2), && (25)\\ R_2 &\leq I(U,V_2;Y_2) - I(V_2;W|U) - I(V_2;Y_1), && (26)\\ R_1+R_2 &\leq I(U,V_1;Y_1) - I(V_1;W|U) + I(U,V_2;Y_2|V_1) \\ &\quad -I(V_2;W|U,V_1) - I(V_1;Y_2), && (27)\\ R_1+R_2 &\leq I(U,V_2;Y_2) - I(V_2;W|U) + I(U,V_1;Y_1|V_2) \\ &\quad -I(V_1;W|U,V_2) - I(V_2;Y_1), && (28) \end{aligned}$$

where $(U, V_1, V_2) \to (X, W) \to (Y_1, Y_2)$.

Theorem 3.8 *Let* $\mathcal{R}_{3,out}^* = \bigcup_{p_3^*(.)\in\mathcal{P}_3^*} \mathcal{R}_{3,out}^*(p_3^*)$. *Then,* $\mathcal{C}_3 \subseteq \mathcal{R}_{3,out}^*$.

The proof of Theorem 3.8 can be found in Section 5.5.

3.4 Discussion

A pictorial representation of the rate region for the channel C_1 is shown in Fig. 2. When $R_2 = 0$, the channel resembles a single-user channel $(\mathrm{S}, \mathrm{D}_1)$ with side-information (the Gel'fand-Pinsker's (GP) channel [20]) and S can transmit at the maximum achievable R_1 given by (3), denoted by point the H. At the point H, the maximum achievable R_2 is given by the point $\mathrm{E}_1 \equiv I(V_2;Y_2) - I(V_1;V_2) - I(W;V_2)$; this is obtained by treating the channel $(\mathrm{S}, \mathrm{D}_2)$ as a single-user channel with side-information. Therefore, the rectangle OHGE_1 is achievable. By exchanging R_1 and R_2 and following similar arguments the points E, given by (4), and $\mathrm{F}_1 \equiv I(V_1;Y_1) - I(V_1;V_2|U) - I(W;V_1)$ are achievable. Hence, the rectangle OEFF_1 is also achievable. Since the points F and G are shown to be achievable, any point which lies on the line FG can also be achieved by deriving a bound on the binning rates (see (65) - (67), Appendix A). This leads to a sum rate bound given by (5). Finally, owing to convexity of the rate region, any point in the interior of the line FG is also achievable. Therefore, an achievable rate region for C_1 is described by the pentagon OEFGH.

In the absence of side-information, *i.e.*, $\mathcal{W} = \{\phi\}$, the channel reduces to the classical two-user BC whose rate region is described by the convex-hull of the set of all rate pairs (R_1, R_2) that satisfy the following inequalities:

$$R_1 \leq I(V_1; Y_1), \tag{29}$$

$$R_2 \leq I(V_2; Y_2), \tag{30}$$

$$R_1 + R_2 \leq I(V_1; Y_1) + I(V_2; Y_2) - I(V_1; V_2). \tag{31}$$

For channels of Class II, each bound in (12) - (13) is the capacity of GP's single-user channel with noncausal side-information. In the absence of side-information, *i.e.*, $\mathcal{W} = \{\phi\}$, we get $R_t \leq I(U; Y_t) = I(X; Y_t)$, which represents the capacity region of BC when each receiver is given the message it need not decode [10]. Furthermore, the outer bounds (14) - (15) is within a fixed gap, $H(U)$, from the achievable region, where $H(U)$ is independent of the distribution characterizing this class of channels.

For Class III channels, the terms $I(V_1; Y_2|U, V_2)$ and $I(V_2; Y_1|U, V_1)$ quantify the rate-penalty for having to deal with confidentiality constraints on the messages, while the terms $I(V_1; W|U)$ and $I(V_2; W|U)$ quantify the rate-penalty for having to deal with side-information.

Using a combination of results from GP's channel and wiretap channels with side-information [21], we obtain a pictorial representation of the rate region for the channel C_3 as shown in Fig. 3. The arguments used to obtain this schematic are similar to those used for the channel C_1; therefore, we briefly explain the construction of Fig. 3. The point A_1 corresponds to the maximum achievable R_1 (when $R_2 = 0$) and is given by (16). Exchanging R_1 and R_2 we get the point C_1 given by (17). The points $\mathrm{B}_1 \equiv I(V_2; Y_2|U) - I(V_2; Y_1|U, V_1) - \max[I(V_1; V_2|U), I(W; V_2|U)]$ and $\mathrm{D}_1 \equiv I(V_1; Y_1|U) - I(V_1; Y_2|U, V_2) - \max[I(V_1; V_2|U), I(W; V_1|U)]$ are achievable by treating channels $(\mathrm{S}, \mathrm{D}_2)$ and $(\mathrm{S}, \mathrm{D}_1)$, respectively, as wiretap channels with side-information. The line $\mathrm{E}_1\mathrm{F}_1$ corresponds to the sum rate bound given by (18). Finally, owing to convexity of the rate region, any point in the interior of the line $\mathrm{E}_1\mathrm{F}_1$ is also achievable. Therefore, an achievable rate region for C_3 is described by the pentagon $\mathrm{OA}_1\mathrm{F}_1\mathrm{E}_1\mathrm{C}_1$.

If the confidentiality constraints (1) - (2) are relaxed, the channel C_3 reduces to the channel C_1, whose rate region is described by (3) - (5). Further, in the absence of side-information, *i.e.*, $\mathcal{W} = \{\phi\}$, the channel reduces to the classical two-user BC whose rate region is described by (29) - (31). Lastly, if the encoder satisfies confidentiality constraints in the absence of side-information, the channel C_3 reduces to BC with two independent and confidential messages whose rate region was first characterized by Liu *et. al* [14]. It is described by the convex-hull of the set of all rate pairs (R_1, R_2) that satisfy the following inequalities:

$$R_1 \leq I(V_1; Y_1|U) - I(V_1; Y_2|U) - I(V_1; V_2|U), \tag{32}$$

$$R_2 \leq I(V_2; Y_2|U) - I(V_2; Y_1|U) - I(V_1; V_2|U). \tag{33}$$

3.5 Relation to past work

For Class I channels, an inner bound was presented in [8] by extending Marton's achievability scheme for the classical two-user BC to include noncausal side-information at the encoder. In this paper, we employ Marton's technique and use results from the second moment method [22] to derive the inner bound which matches with the results presented in [8]. However, our method is simpler and generalizes well for obtaining inner bounds with other channel models, *e.g.*, for channels of Class III considered in this paper. For the outer bound (specifically, for the sum-rate), we generalize the technique presented in [5], to handle side-information at the encoder. When the side-information constraint is relaxed, our result reduces to the one presented for the classical two-user BC [5].

Class II channels were also addressed in [23], where an inner bound was derived by employing Marton's achievability scheme. An outer bound was also suggested in [23], but without a formal proof. In this paper, we derive an inner bound by generalizing the method suggested in [10] by incorporating noncausal side-information at the encoder. Our inner bound coincides with the one presented in [23], but once again the proof technique is much simpler. Furthermore, for the outer bounds, we explicitly address the problem of dealing with the two-dimensional rate region with a single auxiliary random variable.

For Class III channels, we show that when the confidentiality constraints are relaxed, our achievable rate region reduces to region presented for the Class I channels, and hence to the one presented in [8]. On the other hand, in the absence of side-information, our achievable region includes an explicit bound on the sum-rate for the two-user BC with confidentiality constraints (a model considered in [14]). This further strengthens the generalization of our proof technique.

4 Proofs of achievability theorems

In this section, we prove Theorem 3.1, Theorem 3.4 and Theorem 3.6. For any $\epsilon > 0$, we denote by $A_{\epsilon}^{(\mathrm{N})}(P_{X})$ an ϵ-typical set comprising sequences picked from the distribution $p(\mathbf{x})$. For all the channel models, the encoder is given an $\epsilon-$typical sequence $\mathbf{W} \in A_{\epsilon}^{(\mathrm{N})}(P_{W})$ in a noncausal manner.

4.1 Proof of Theorem 3.1

For the channel C_1, generate $2^{\mathrm{N}[R_t+R_t^{'}]}$ independent typical sequences $\mathbf{V}_t(i_t, j_t) \in A_{\epsilon}^{(\mathrm{N})}(P_{V_t}); t = 1, 2$. Here, $i_t \in \{1, \ldots, 2^{\mathrm{N}R_t}\}$; $j_t \in \{1, \ldots, 2^{\mathrm{N}R_t^{'}}\}$. Uniformly distribute $2^{\mathrm{N}[R_t+R_t^{'}]}$ sequences into $2^{\mathrm{N}R_t}$ bins, so that each bin, indexed by i_t, comprises $2^{\mathrm{N}R_t^{'}}$ sequences. To send the message pair $(m_1 = i_1, m_2 = i_2)$, the encoder at S looks for a pair (j_1, j_2) that satisfies the following joint typicality condition: $E_{\mathrm{S}} \triangleq \{(\mathbf{W}, \mathbf{V}_1(i_1, j_1), \mathbf{V}_2(i_2, j_2)) \in A_{\epsilon}^{(\mathrm{N})}(P_{W,V_1,V_2})\}$. An error is declared at the encoder of S, if it is not possible to find the $(j_1, j_2)-$pair to satisfy the condition E_{S}. The encoder error analysis can be found in Appendix A. The channel input sequence is $\mathbf{X} \in A_{\epsilon}^{(\mathrm{N})}(P_{X|W,V_1,V_2})$.

At the destination D_t, the decoder looks for $(\hat{i}_t, \hat{j}_t)$ that satisfies the following joint typicality condition: $E_{\mathrm{D}_t} \triangleq \{(\mathbf{V}_t(\hat{i}_t, \hat{j}_t), \mathbf{Y}_t) \in A_{\epsilon}^{(\mathrm{N})}(P_{V_t,Y_t})\}$. An error is declared at decoder of D_t, if it not possible to find a unique integer $\hat{i}_t$ to satisfy the condition E_{D_t}. From the union of events bound, the probability of decoder error at D_t can be upper bounded as follows: $P_{e,\mathrm{D}_t}^{(\mathrm{N})} \leq \Pr(E_{\mathrm{D}_t}^{c}|E_{\mathrm{S}}) + \sum_{\hat{i}_t \neq i_t} \sum_{j_t} \Pr(E_{\mathrm{D}_t}|E_{\mathrm{S}})$. From the asymptotic equipartition property (AEP) [24], $\forall\epsilon > 0$ and sufficiently small; and for large N, $\Pr(E_{\mathrm{D}_t}^{c}|E_{\mathrm{S}}) \leq \epsilon$. Further, for $\hat{i}_t \neq i_t$, $\Pr(E_{\mathrm{D}_t}|E_{\mathrm{S}}) \leq 2^{-\mathrm{N}[I(V_t;Y_t)-\epsilon]}$. Therefore, we have $P_{e,\mathrm{D}_t}^{(\mathrm{N})} \leq \epsilon + 2^{\mathrm{N}[R_t+R_t^{'}]}2^{-\mathrm{N}[I(V_t;Y_t)-\epsilon]}$, leading us to conclude that, for any $\epsilon_0 > 0$ and sufficiently small; and for large N, $P_{e,\mathrm{D}_t}^{(\mathrm{N})} \leq \epsilon_0$ if

$$R_t + R_t^{'} < I(V_t; Y_t). \tag{34}$$

For the channel C_1, the rate inequalities (34) and the bounds on the binning rates (65) - (67) (see Appendix A) are combined to obtain an achievable rate region given by (3) - (5). This completes the proof of Theorem 3.1.

4.2 Proof of Theorem 3.4

For the channel C_2, we consider the following two cases.

1. When $R_1 \leq R_2$: Generate $2^{\mathrm{N}(R_2+R^*)}$ typical sequences $\mathbf{U}(i,j) \in A_\epsilon^{(\mathrm{N})}(P_U); i \in \{1,\ldots,2^{\mathrm{N}R_2}\}$; $j \in \{1,\ldots,2^{\mathrm{N}R^*}\}$. Uniformly distribute these sequences into $2^{\mathrm{N}R_2}$ bins, so that each bin comprises $2^{\mathrm{N}R^*}$ sequences. The bins are indexed by i. Define now the following mappings:

$$m_t \in \{1,\ldots,2^{\mathrm{N}R_t}\} \quad \mapsto \quad \mathrm{Int}(m_t) \in \{0,\ldots,2^{\mathrm{N}R_2}-1\}; t = 1,2,$$

   where $\mathrm{Int}(\alpha)$ denotes an integer to represent α. To transmit the message pair (m_1, m_2), compute $\left(\mathrm{Int}(m_1) + \mathrm{Int}(m_2) \mod 2^{\mathrm{N}R_2}\right)$. By construction, the bin index $i \triangleq \mathrm{Int}^{-1}\left(\mathrm{Int}(m_1) + \mathrm{Int}(m_2) \mod 2^{\mathrm{N}R_2}\right)$. Given the sequence $\mathbf{W}$, the encoder looks for an integer j to satisfy the following joint typicality condition:

$$(\mathbf{U}(i,j), \mathbf{W}) \in A_\epsilon^{(\mathrm{N})}(P_{W,U}).$$

   Finally, $\mathbf{X} \triangleq \mathbf{f}(\mathbf{U}(i,j), \mathbf{W})$ is transmitted in N channel uses.

   At receiver D_1, given m_2, the decoder looks for the pair $(\hat{\imath} \triangleq \hat{m}_1, \hat{\jmath})$ such that the following joint typicality condition is satisfied:

$$E_{\mathrm{D}_1} \triangleq \{(\mathbf{U}(\mathrm{Int}^{-1}\left(\mathrm{Int}(\hat{m}_1) + \mathrm{Int}(m_2) \mod 2^{\mathrm{N}R_2}\right), j), \mathbf{Y}_1) \in A_\epsilon^{(\mathrm{N})}(P_{U,Y_1})\}.$$

   From AEP, it can be shown that $\Pr(E_{\mathrm{D}_1}^c) \leq \delta_1; \forall \delta_1 > 0$ and sufficiently small; and for large N, if $R_1 + R^* \leq I(U;Y_1)$. Similarly, it can be shown that $\Pr(E_{\mathrm{D}_2}^c) \leq \delta_2; \forall \delta_2 > 0$ and sufficiently small; and for large N, if $R_2 + R^* \leq I(U;Y_2)$. Additionally, by following a procedure similar to the one presented in Appendix A, we bound the binning rate as follows: $R^* > I(U;W)$. Therefore, m_1 (resp. m_2) can be reliably decoded at D_1 (resp. D_2) if

$$R_1 \leq I(U;Y_1) - I(U;W), \tag{35}$$

$$R_2 \leq I(U;Y_2) - I(U;W). \tag{36}$$

2. When $R_2 \leq R_1$: By symmetry, we get the same rate bounds as in (35) and (36).

This completes the proof of Theorem 3.4.

4.3 Proof of Theorem 3.6

For the channel C_3, generate a typical sequence $\mathbf{U} \in A_\epsilon^{(\mathrm{N})}(P_U)$, known to all nodes in the network. Generate $2^{\mathrm{N}[R_t+R_t^{'}+R_t^*]}$ independent typical sequences $\mathbf{V}_t(i_t, j_t, k_t) \in A_\epsilon^{(\mathrm{N})}(P_{V_t})$; $i_t \in \{1,\ldots,2^{\mathrm{N}R_t}\}$;

$j_t \in \{1, \ldots, 2^{\mathrm{N}R'_t}\}$; $k_t \in \{1, \ldots, 2^{\mathrm{N}R^*_t}\}$. Uniformly distribute $2^{\mathrm{N}[R_t+R'_t+R^*_t]}$ sequences into $2^{\mathrm{N}R_t}$ bins, so that each bin, indexed by i_t, comprises $2^{\mathrm{N}[R'_t+R^*_t]}$ sequences. Uniformly distribute $2^{\mathrm{N}[R'_t+R^*_t]}$ sequences into $2^{\mathrm{N}R'_t}$ sub-bins indexed by (i_t, j_t), so that each sub-bin comprises $2^{\mathrm{N}R^*_t}$ sequences.

To send the message pair $(\mathsf{m}_1, \mathsf{m}_2)$, S employs a stochastic encoder. In the bin indexed by i_t, *randomly* pick a sub-bin indexed (i_t, j_t). The encoder then looks for a pair (k_1, k_2) that satisfies the following joint typicality condition: $(\mathbf{W}, \mathbf{V}_1(i_1, j_1, k_1), \mathbf{V}_2(i_2, j_2, k_2)) \in A_\epsilon^{(\mathrm{N})}(P_{W,V_1,V_2|U})$. The channel input sequence $\mathbf{X} \in A_\epsilon^{(\mathrm{N})}(P_{X|W,V_1,V_2})$ is transmitted in N uses of the channel.

At the destination D_t, given $\mathbf{U}$, the decoder picks k_t that satisfies the following joint typicality condition: $E_{\mathrm{D}_t} \triangleq \{(\mathbf{V}_t(i_t, j_t, k_t), \mathbf{Y}_t) \in A_\epsilon^{(\mathrm{N})}(P_{V_t,Y_t|U})\}$. An error is declared at the decoder of D_t if it not possible to find an integer $\hat{i}_t$ satisfying E_{D_t}. From union of events bound, the probability of decoder error at D_t can be upper bounded as follows: $P_{e,\mathrm{D}_t}^{(\mathrm{N})} \leq \Pr(E^c_{\mathrm{D}_t}|E_{\mathrm{S}}) + \sum_{\hat{i}_t \neq i_t} \sum_{j_t,k_t} \Pr(E_{\mathrm{D}_t}|E_{\mathrm{S}})$. From AEP [24], $\forall \epsilon > 0$ and sufficiently small; and for large N, $\Pr(E^c_{\mathrm{D}_t}|E_{\mathrm{S}}) \leq \epsilon$ and for $\hat{i}_t \neq i_t$, we have $\Pr(E_{\mathrm{D}_t}|E_{\mathrm{S}}) \leq 2^{-\mathrm{N}[I(V_t;Y_t|U)-\epsilon]}$. Therefore, $P_{e,\mathrm{D}_t}^{(\mathrm{N})} \leq \epsilon + 2^{\mathrm{N}[R_t+R'_t+R^*_t]} 2^{-\mathrm{N}[I(V_t;Y_t|U)-\epsilon]}$. For any $\epsilon_0 > 0$ and sufficiently small; and for large N, $P_{e,\mathrm{D}_t}^{(\mathrm{N})} \leq \epsilon_0$ if

$$R_t + R'_t + R^*_t < I(V_t; Y_t|U). \tag{37}$$

The equivocation at the decoder of D_2 is calculated by first considering the following lower bound: $H(M_1|\mathbf{Y}_2^{\mathrm{N}}) \geq H(M_1|\mathbf{Y}_2^{\mathrm{N}}, \mathbf{U}^{\mathrm{N}}, \mathbf{V}_2^{\mathrm{N}})$. Following the procedure in [14, Section V-B] and using the fact that $M_1 \to (\mathbf{U}^{\mathrm{N}}, \mathbf{V}_1^{\mathrm{N}}, \mathbf{V}_2^{\mathrm{N}}) \to \mathbf{Y}_2^{\mathrm{N}}$ forms a Markov chain, we get

$$H(M_1|\mathbf{Y}_2^{\mathrm{N}}) \geq H(\mathbf{V}_1^{\mathrm{N}}|\mathbf{U}^{\mathrm{N}}) - I(\mathbf{V}_1^{\mathrm{N}}; \mathbf{V}_2^{\mathrm{N}}|\mathbf{U}^{\mathrm{N}}) - H(\mathbf{V}_1^{\mathrm{N}}|M_1, \mathbf{U}^{\mathrm{N}}, \mathbf{V}_2^{\mathrm{N}}, \mathbf{Y}_2^{\mathrm{N}}) - I(\mathbf{V}_1^{\mathrm{N}}; \mathbf{Y}_2^{\mathrm{N}}|\mathbf{U}^{\mathrm{N}}, \mathbf{V}_2^{\mathrm{N}}). \tag{38}$$

Now, $\forall \epsilon_l > 0; l = 4, \ldots, 10$ and sufficiently small; and for large N, the terms in (38) become

$$\begin{gathered} H(\mathbf{V}_1^{\mathrm{N}}|\mathbf{U}^{\mathrm{N}}) \overset{(a)}{=} \mathrm{N}[R_1 + R'_1 + R^*_1]; I(\mathbf{V}_1^{\mathrm{N}}; \mathbf{V}_2^{\mathrm{N}}|\mathbf{U}^{\mathrm{N}}) \overset{(b)}{=} \mathrm{N}I(V_1; V_2|U) + \mathrm{N}\epsilon_4; \\ H(\mathbf{V}_1^{\mathrm{N}}|M_1, \mathbf{U}^{\mathrm{N}}, \mathbf{V}_2^{\mathrm{N}}, \mathbf{Y}_2^{\mathrm{N}}) \overset{(c)}{\leq} \mathrm{N}\epsilon_5; I(\mathbf{V}_1^{\mathrm{N}}; \mathbf{Y}_2^{\mathrm{N}}|\mathbf{U}^{\mathrm{N}}, \mathbf{V}_2^{\mathrm{N}}) \overset{(d)}{=} \mathrm{N}I(V_1; Y_2|U, V_2) + \mathrm{N}\epsilon_6. \end{gathered} \tag{39}$$

In (39), (a) follows from the codebook construction; (b) and (d) follow from standard techniques (for *e.g.*, see [14, Lemma 3]); and (c) is proved in [14, Lemma 2]. A similar procedure is followed to calculate the equivocation at the decoder at D_1. Finally, the security constraints (1) and (2) are satisfied by letting

$$R'_1 = I(V_1; Y_2|U, V_2) - \epsilon_7; R^*_1 = I(V_1; V_2|U) - \epsilon_8; \tag{40}$$

$$R'_2 = I(V_2; Y_1|W, U, V_1) - \epsilon_9; R^*_2 = I(V_1; V_2|W, U) - \epsilon_{10}. \tag{41}$$

For the channel C_3, rate inequalities (37), constraints (40) - (41) and bounds on the binning rates (68) - (70) (see Appendix A) are combined to obtain the rate region described by (16) - (18). This completes the proof of Theorem 3.6.

5 Proofs of converse theorems

In this section, we prove Theorem 3.2, Theorem 3.3, Theorem 3.5, Theorem 3.7 and Theorem 3.8.

5.1 Proof of Theorem 3.2

For the channel C_1, $\forall\epsilon > 0$ and sufficiently small; and for large N, R_1 can be bounded as follows:

$$\begin{aligned}
\mathrm{N}R_1 &= H(M_1) = I(M_1;\mathbf{Y}_1^{\mathrm{N}}) + H(M_1|\mathbf{Y}_1^{\mathrm{N}}) \\
&\overset{(a)}{\leq} I(M_1;\mathbf{Y}_1^{\mathrm{N}}) + \mathrm{N}\epsilon \overset{(b)}{=} \sum_{\mathrm{n}=1}^{\mathrm{N}}[H(Y_{1,\mathrm{n}}|\mathbf{Y}_1^{\mathrm{n}-1}) - H(Y_{1,\mathrm{n}}|\mathbf{Y}_1^{\mathrm{n}-1}, M_1)] + \mathrm{N}\epsilon \\
&\overset{(c)}{\leq} \sum_{\mathrm{n}=1}^{\mathrm{N}}[H(Y_{1,\mathrm{n}}) - H(Y_{1,\mathrm{n}}|\mathbf{Y}_1^{\mathrm{n}-1}, M_1)] + \mathrm{N}\epsilon = \sum_{\mathrm{n}=1}^{\mathrm{N}} I(M_1,\mathbf{Y}_1^{\mathrm{n}-1}; Y_{1,\mathrm{n}}) + \mathrm{N}\epsilon \\
&= \sum_{\mathrm{n}=1}^{\mathrm{N}}[I(M_1,\mathbf{Y}_1^{\mathrm{n}-1},\mathbf{W}_{\mathrm{n}+1}^{\mathrm{N}}; Y_{1,\mathrm{n}}) - I(\mathbf{W}_{\mathrm{n}+1}^{\mathrm{N}}; Y_{1,\mathrm{n}}|M_1,\mathbf{Y}_1^{\mathrm{n}-1})] + \mathrm{N}\epsilon \\
&\overset{(d)}{=} \sum_{\mathrm{n}=1}^{\mathrm{N}}[I(M_1,\mathbf{Y}_1^{\mathrm{n}-1},\mathbf{W}_{\mathrm{n}+1}^{\mathrm{N}}; Y_{1,\mathrm{n}}) - I(\mathbf{Y}_1^{\mathrm{n}-1}; W_{\mathrm{n}}|M_1,\mathbf{W}_{\mathrm{n}+1}^{\mathrm{N}})] + \mathrm{N}\epsilon \\
&\overset{(e)}{=} \sum_{\mathrm{n}=1}^{\mathrm{N}}[I(M_1,\mathbf{Y}_1^{\mathrm{n}-1},\mathbf{W}_{\mathrm{n}+1}^{\mathrm{N}}; Y_{1,\mathrm{n}}) - I(M_1,\mathbf{W}_{\mathrm{n}+1}^{\mathrm{N}},\mathbf{Y}_1^{\mathrm{n}-1}; W_{\mathrm{n}})] + \mathrm{N}\epsilon,
\end{aligned}$$

where (a) follows from Fano's inequality [24], (b) follows from the chain rule, (c) follows from the fact that conditioning reduces entropy, (d) follows from Csiszár's sum identity [25] and (e) is due to the fact that $(M_1,\mathbf{W}_{\mathrm{n}+1}^{\mathrm{N}})$ is independent of W_{n}. We let $V_{1,\mathrm{n}} = (M_1,\mathbf{W}_{\mathrm{n}+1}^{\mathrm{N}},\mathbf{Y}_1^{\mathrm{n}-1})$ and note that this choice satisfies the Markov chain requirement $V_1 \to (X, W) \to (Y_1, Y_2)$, specified in Section 3 for the channel C_1. Thus, we get

$$\mathrm{N}R_1 \leq \sum_{\mathrm{n}=1}^{\mathrm{N}} I(V_{1,\mathrm{n}}; Y_{1,\mathrm{n}}) - I(V_{1,\mathrm{n}}; W_{\mathrm{n}}) + \mathrm{N}\epsilon. \tag{42}$$

Proceeding in a similar manner and letting $V_{2,\mathrm{n}} = (M_2,\mathbf{W}_{\mathrm{n}+1}^{\mathrm{N}},\mathbf{Y}_2^{\mathrm{n}-1})$, we get

$$\mathrm{N}R_2 \leq \sum_{\mathrm{n}=1}^{\mathrm{N}} I(V_{2,\mathrm{n}}; Y_{2,\mathrm{n}}) - I(V_{2,\mathrm{n}}; W_{\mathrm{n}}) + \mathrm{N}\epsilon. \tag{43}$$

5.2 Proof of Theorem 3.3

For the channel C_1, $\forall\epsilon > 0$ and sufficiently small; and for large N, R_1 can be bounded as

$$\begin{aligned}
\mathrm{N}R_1 &= H(M_1) = I(M_1;\mathbf{Y}_1^{\mathrm{N}}) + H(M_1|\mathbf{Y}_1^{\mathrm{N}})\\
&\overset{(a)}{\leq} I(M_1;\mathbf{Y}_1^{\mathrm{N}}) + \mathrm{N}\epsilon,
\end{aligned}$$

where (a) follows from Fano's inequality. Proceeding in a manner similar to the proof of Theorem 3.2 (see Section 5.1), and letting $U_{\mathrm{n}} = (\mathbf{W}_{\mathrm{n}+1}^{\mathrm{N}}, \mathbf{Y}_1^{\mathrm{n}-1}, \mathbf{Y}_{2,\mathrm{n}+1}^{\mathrm{N}})$ and $V_{1,\mathrm{n}} = M_1$.

$$\mathrm{N}R_1 \leq \sum_{\mathrm{n}=1}^{\mathrm{N}} I(U_{\mathrm{n}}, V_{1,\mathrm{n}}; Y_{1,\mathrm{n}}) - I(V_{1,\mathrm{n}}; W_{\mathrm{n}}|U_{\mathrm{n}}) + \mathrm{N}\epsilon. \tag{44}$$

Similarly, letting $V_{2,\mathrm{n}} = M_2$, R_2 can be upper bounded as follows:

$$\mathrm{N}R_2 \leq \sum_{\mathrm{n}=1}^{\mathrm{N}} I(U_{\mathrm{n}}, V_{2,\mathrm{n}}; Y_{2,\mathrm{n}}) - I(V_{2,\mathrm{n}}; W_{\mathrm{n}}|U_{\mathrm{n}}) + \mathrm{N}\epsilon. \tag{45}$$

We next upper bound $R_1 + R_2$ as follows. $\forall\epsilon > 0$ and sufficiently small; and for large N, we have

$$\begin{aligned}
\mathrm{N}(R_1 + R_2) &= H(M_1, M_2) = H(M_1) + H(M_2|M_1)\\
&= I(M_1;\mathbf{Y}_1^{\mathrm{N}}) + H(M_1|\mathbf{Y}_1^{\mathrm{N}}) + I(M_2;\mathbf{Y}_2^{\mathrm{N}}|M_1) + H(M_2|\mathbf{Y}_2^{\mathrm{N}}, M_1)\\
&\overset{(a)}{\leq} \sum_{\mathrm{n}=1}^{\mathrm{N}} I(M_1; Y_{1,\mathrm{n}}|\mathbf{Y}_1^{\mathrm{n}-1}) + \sum_{\mathrm{n}=1}^{\mathrm{N}} I(M_2; Y_{2,\mathrm{n}}|\mathbf{Y}_{2,\mathrm{n}+1}^{\mathrm{N}}, M_1) + 2\mathrm{N}\epsilon,
\end{aligned}$$

where (a) follows from Fano's inequality. Consider

$$\begin{aligned}
\sum_{\mathrm{n}=1}^{\mathrm{N}} I(M_1;\mathbf{Y}_{1,\mathrm{n}}|\mathbf{Y}_1^{\mathrm{n}-1}) &\leq \sum_{\mathrm{n}=1}^{\mathrm{N}} I(M_1, \mathbf{Y}_1^{\mathrm{n}-1};\mathbf{Y}_{1,\mathrm{n}})\\
&= \sum_{\mathrm{n}=1}^{\mathrm{N}} I(M_1, \mathbf{Y}_1^{\mathrm{n}-1}, \mathbf{Y}_{2,\mathrm{n}+1}^{\mathrm{N}};\mathbf{Y}_{1,\mathrm{n}}) - \sum_{\mathrm{n}=1}^{\mathrm{N}} I(\mathbf{Y}_{2,\mathrm{n}+1}^{\mathrm{N}};\mathbf{Y}_{1,\mathrm{n}}|M_1, \mathbf{Y}_1^{\mathrm{n}-1})\\
&= \sum_{\mathrm{n}=1}^{\mathrm{N}} [I(M_1, \mathbf{Y}_1^{\mathrm{n}-1}, \mathbf{Y}_{2,\mathrm{n}+1}^{\mathrm{N}}, \mathbf{W}_{\mathrm{n}+1}^{\mathrm{N}}; Y_{1,\mathrm{n}}) - I(\mathbf{W}_{\mathrm{n}+1}^{\mathrm{N}}; Y_{1,\mathrm{n}}|M_1, \mathbf{Y}_1^{\mathrm{n}-1}, \mathbf{Y}_{2,\mathrm{n}+1}^{\mathrm{N}})]\\
&- \sum_{\mathrm{n}=1}^{\mathrm{N}} I(\mathbf{Y}_{2,\mathrm{n}+1}^{\mathrm{N}}; Y_{1,\mathrm{n}}|M_1, \mathbf{Y}_1^{\mathrm{n}-1})\\
&\overset{(b)}{=} \sum_{\mathrm{n}=1}^{\mathrm{N}} [I(M_1, \mathbf{Y}_1^{\mathrm{n}-1}, \mathbf{Y}_{2,\mathrm{n}+1}^{\mathrm{N}}, \mathbf{W}_{\mathrm{n}+1}^{\mathrm{N}}; Y_{1,\mathrm{n}}) - I(M_1; W_{\mathrm{n}}|\mathbf{W}_{\mathrm{n}+1}^{\mathrm{N}}, \mathbf{Y}_1^{\mathrm{n}-1}, \mathbf{Y}_{2,\mathrm{n}+1}^{\mathrm{N}})]\\
&- \sum_{\mathrm{n}=1}^{\mathrm{N}} I(\mathbf{Y}_{2,\mathrm{n}+1}^{\mathrm{N}}; Y_{1,\mathrm{n}}|M_1, \mathbf{Y}_1^{\mathrm{n}-1})
\end{aligned} \tag{46}$$

Next consider

$$\begin{aligned}\sum_{\mathrm{n}=1}^{\mathrm{N}} I(M_2; Y_{2,\mathrm{n}}|\mathbf{Y}_{2,\mathrm{n}+1}^{\mathrm{N}}, M_1) &\leq \sum_{\mathrm{n}=1}^{\mathrm{N}} I(M_2, \mathbf{Y}_1^{\mathrm{n}-1}; Y_{2,\mathrm{n}}|\mathbf{Y}_{2,\mathrm{n}+1}^{\mathrm{N}}, M_1) \\
= \sum_{\mathrm{n}=1}^{\mathrm{N}} I(\mathbf{Y}_1^{\mathrm{n}-1}; Y_{2,\mathrm{n}}|\mathbf{Y}_{2,\mathrm{n}+1}^{\mathrm{N}}, M_1) &+ \sum_{\mathrm{n}=1}^{\mathrm{N}} I(M_2; \mathbf{Y}_{2,\mathrm{n}}|\mathbf{Y}_1^{\mathrm{n}-1}, \mathbf{Y}_{2,\mathrm{n}+1}^{\mathrm{N}}, M_1) \\
= \sum_{\mathrm{n}=1}^{\mathrm{N}} I(\mathbf{Y}_1^{\mathrm{n}-1}; Y_{2,\mathrm{n}}|\mathbf{Y}_{2,\mathrm{n}+1}^{\mathrm{N}}, M_1) &+ \sum_{\mathrm{n}=1}^{\mathrm{N}} I(M_2, \mathbf{W}_{\mathrm{n}+1}^{\mathrm{N}}; Y_{2,\mathrm{n}}|\mathbf{Y}_1^{\mathrm{n}-1}, \mathbf{Y}_{2,\mathrm{n}+1}^{\mathrm{N}}, M_1) \\
&- \sum_{\mathrm{n}=1}^{\mathrm{N}} I(\mathbf{W}_{\mathrm{n}+1}^{\mathrm{N}}; Y_{2,\mathrm{n}}|\mathbf{Y}_1^{\mathrm{n}-1}, \mathbf{Y}_{2,\mathrm{n}+1}^{\mathrm{N}}, M_1, M_2) \\
= \sum_{\mathrm{n}=1}^{\mathrm{N}} I(\mathbf{Y}_1^{\mathrm{n}-1}; Y_{2,\mathrm{n}}|\mathbf{Y}_{2,\mathrm{n}+1}^{\mathrm{N}}, M_1) &+ \sum_{\mathrm{n}=1}^{\mathrm{N}} I(M_2, \mathbf{Y}_1^{\mathrm{n}-1}, \mathbf{Y}_{2,\mathrm{n}+1}^{\mathrm{N}}, \mathbf{W}_{\mathrm{n}+1}^{\mathrm{N}}; Y_{2,n}|M_1) \\
&- \sum_{\mathrm{n}=1}^{\mathrm{N}} I(M_2; W_{\mathrm{n}}|\mathbf{W}_{\mathrm{n}+1}^{\mathrm{N}}, \mathbf{Y}_1^{\mathrm{n}-1}, \mathbf{Y}_{2,\mathrm{n}+1}^{\mathrm{N}}, M_1) \\
\overset{(c)}{=} \sum_{\mathrm{n}=1}^{\mathrm{N}} I(\mathbf{Y}_1^{\mathrm{n}-1}; Y_{2,\mathrm{n}}|\mathbf{Y}_{2,\mathrm{n}+1}^{\mathrm{N}}, M_1) &+ \sum_{\mathrm{n}=1}^{\mathrm{N}} I(M_2, \mathbf{Y}_1^{\mathrm{n}-1}, \mathbf{Y}_{2,\mathrm{n}+1}^{\mathrm{N}}, \mathbf{W}_{\mathrm{n}+1}^{\mathrm{N}}; Y_{2,\mathrm{n}}|M_1) \\
&- \sum_{\mathrm{n}=1}^{\mathrm{N}} I(M_2; W_{\mathrm{n}}|\mathbf{W}_{\mathrm{n}+1}^{\mathrm{N}}, \mathbf{Y}_1^{\mathrm{n}-1}, \mathbf{Y}_{2,\mathrm{n}+1}^{\mathrm{N}}, M_1) \qquad (47)\end{aligned}$$

where (b) and (c) follow from Csiszár's sum identity. With $U_{\mathrm{n}} = (\mathbf{W}_{\mathrm{n}+1}^{\mathrm{N}}, \mathbf{Y}_1^{\mathrm{n}-1}, \mathbf{Y}_{2,\mathrm{n}+1}^{\mathrm{N}})$; $V_{1,\mathrm{n}} = M_1$; and $V_{2,\mathrm{n}} = M_2$, from (46) and (47), we get

$$\begin{aligned}\mathrm{N}(R_1 + R_2) &\leq \sum_{n=1}^{\mathrm{N}} [I(U_{\mathrm{n}}, V_{1,\mathrm{n}}; Y_{1,\mathrm{n}}) - I(V_{1,\mathrm{n}}; W_{\mathrm{n}}|U_{\mathrm{n}})] \\
&+ \sum_{\mathrm{n}=1}^{\mathrm{N}} [I(U_{\mathrm{n}}, V_{2,\mathrm{n}}; Y_{2,\mathrm{n}}|V_{1,\mathrm{n}}) - I(V_{2,\mathrm{n}}; W_{\mathrm{n}}|V_{1,\mathrm{n}}, U_{\mathrm{n}})] + 2\mathrm{N}\epsilon. \qquad (48)\end{aligned}$$

Similarly, it can be shown that

$$\begin{aligned}\mathrm{N}(R_1 + R_2) &\leq \sum_{\mathrm{n}=1}^{\mathrm{N}} [I(U_{\mathrm{n}}, V_{2,\mathrm{n}}; Y_{2,\mathrm{n}}) - I(V_{2,\mathrm{n}}; W_{\mathrm{n}}|U_{\mathrm{n}})] \\
&+ \sum_{\mathrm{n}=1}^{\mathrm{N}} [I(U_{\mathrm{n}}, V_{1,\mathrm{n}}; Y_{1,\mathrm{n}}|V_{2,\mathrm{n}}) - I(V_{1,\mathrm{n}}; W_{\mathrm{n}}|V_{2,\mathrm{n}}, U_{\mathrm{n}})] + 2\mathrm{N}\epsilon. \qquad (49)\end{aligned}$$

5.3 Proof of Theorem 3.5

For the channel C_2, $\forall\epsilon > 0$ and sufficiently small; and for large N, R_1 can be bounded as follows:

$$\mathrm{N}R_1 \quad = \quad H(M_1) = I(M_1; \mathbf{Y}_1^{\mathrm{N}}) + H(M_1|\mathbf{Y}_1^{\mathrm{N}})$$

$$
\begin{aligned}
&\overset{(a)}{\leq} I(\mathsf{M}_1;\mathbf{Y}_1^{\mathrm{N}})+\mathrm{N}\epsilon \overset{(b)}{\leq} I(\mathsf{M}_1;\mathbf{Y}_1^{\mathrm{N}},\mathsf{M}_2)+\mathrm{N}\epsilon = I(\mathsf{M}_1;\mathbf{Y}_1^{\mathrm{N}}|\mathsf{M}_2)+\mathrm{N}\epsilon\\
&\overset{(c)}{=} \sum_{\mathrm{n}=1}^{\mathrm{N}}[H(\mathsf{Y}_{1,\mathrm{n}}|\mathbf{Y}_1^{\mathrm{n}-1},\mathsf{M}_2)-H(\mathsf{Y}_{1,\mathrm{n}}|\mathbf{Y}_1^{\mathrm{n}-1},\mathsf{M}_1,\mathsf{M}_2)]+\mathrm{N}\epsilon\\
&\overset{(d)}{\leq} \sum_{\mathrm{n}=1}^{\mathrm{N}}[H(\mathsf{Y}_{1,\mathrm{n}})-H(\mathsf{Y}_{1,\mathrm{n}}|\mathbf{Y}_1^{\mathrm{n}-1},\mathsf{M}_1,\mathsf{M}_2)]+\mathrm{N}\epsilon\\
&= \sum_{\mathrm{n}=1}^{\mathrm{N}} I(\mathsf{M}_1,\mathsf{M}_2,\mathbf{Y}_1^{\mathrm{n}-1};\mathsf{Y}_{1,\mathrm{n}})+\mathrm{N}\epsilon\\
&= \sum_{\mathrm{n}=1}^{\mathrm{N}}[I(\mathsf{M}_1,\mathsf{M}_2,\mathbf{Y}_1^{\mathrm{n}-1},\mathbf{W}_{\mathrm{n}+1}^{\mathrm{N}};\mathsf{Y}_{1,\mathrm{n}})-I(\mathbf{W}_{\mathrm{n}+1}^{\mathrm{N}};\mathsf{Y}_{1,\mathrm{n}}|\mathsf{M}_1,\mathsf{M}_2,\mathbf{Y}_1^{\mathrm{n}-1})]+\mathrm{N}\epsilon\\
&\overset{(e)}{=} \sum_{\mathrm{n}=1}^{\mathrm{N}}[I(\mathsf{M}_1,\mathsf{M}_2,\mathbf{Y}_1^{\mathrm{n}-1},\mathbf{W}_{\mathrm{n}+1}^{\mathrm{N}};\mathsf{Y}_{1,\mathrm{n}})-I(\mathsf{M}_1;\mathsf{W}_{\mathrm{n}}|\mathsf{M}_2,\mathbf{Y}_1^{\mathrm{n}-1},\mathbf{W}_{\mathrm{n}+1}^{\mathrm{N}})]+\mathrm{N}\epsilon\\
&\overset{(f)}{=} \sum_{\mathrm{n}=1}^{\mathrm{N}}[I(\mathsf{M}_1,\mathsf{M}_2,\mathbf{W}_{\mathrm{n}+1}^{\mathrm{N}};\mathsf{Y}_{1,\mathrm{n}})-I(\mathsf{M}_1,\mathsf{M}_2,\mathbf{W}_{\mathrm{n}+1}^{\mathrm{N}};\mathsf{W}_{\mathrm{n}}|\mathbf{Y}_1^{\mathrm{n}-1})]+\mathrm{N}\epsilon\\
&\overset{(g)}{\leq} \sum_{\mathrm{n}=1}^{\mathrm{N}}[I(\mathsf{M}_1,\mathsf{M}_2,\mathbf{W}_{\mathrm{n}+1}^{\mathrm{N}};\mathsf{Y}_{1,\mathrm{n}})-I(\mathsf{M}_1,\mathsf{M}_2,\mathbf{W}_{\mathrm{n}+1}^{\mathrm{N}};\mathsf{W}_{\mathrm{n}})+H(\mathsf{M}_1,\mathsf{M}_2,\mathbf{W}_{\mathrm{n}+1}^{\mathrm{N}})]+\mathrm{N}\epsilon. \quad (50)
\end{aligned}
$$

where (a) follows from Fano's inequality; (b) follows from the data-processing inequality; (c) follows from chain rule; (d) follows from the fact that conditioning reduces entropy; (e) follows from Csiszár's sum identity; (f) is due to the memoryless nature of the channel; and (g) is obtained after simple calculations. We let $\mathsf{U}_n \triangleq (\mathsf{M}_1,\mathsf{M}_2,\mathbf{W}_{n+1}^{\mathrm{N}})$ and note that this choice satisfies the Markov chain requirement $\mathsf{U}\rightarrow(\mathsf{X},\mathsf{W})\rightarrow(\mathsf{Y}_1,\mathsf{Y}_2)$ specified in Section 3 for the channel C_2 to get

$$\mathrm{N}R_1 \leq \sum_{n=1}^{\mathrm{N}}[I(\mathsf{U}_n;\mathsf{Y}_{1,n})-I(\mathsf{U}_n;\mathsf{W}_n)+H(\mathsf{U}_n)]+\mathrm{N}\epsilon. \quad (51)$$

By symmetry, we get the following bound on R_2:

$$\mathrm{N}R_2 \leq \sum_{n=1}^{\mathrm{N}}[I(\mathsf{U}_n;\mathsf{Y}_{2,n})-I(\mathsf{U}_n;\mathsf{W}_n)+H(\mathsf{U}_n)]+\mathrm{N}\epsilon. \quad (52)$$

We note that the factor $H(\mathsf{U}_n)$ is independent of the distribution characterizing the channel C_2.

5.4 Proof of Theorem 3.7

For the channel C_3, $\forall\epsilon>0$ and sufficiently small; and for large N, R_1 can be bounded as follows:

$$\mathrm{N}R_1 \quad = \quad H(\mathsf{M}_1)=I(\mathsf{M}_1;\mathbf{Y}_1^{\mathrm{N}})+H(\mathsf{M}_1|\mathbf{Y}_1^{\mathrm{N}})$$

$$
\begin{aligned}
&\overset{(a)}{\leq} \; I(M_1;\mathbf{Y}_1^{\mathrm{N}}) + \mathrm{N}\epsilon \overset{(b)}{\leq} I(M_1;\mathbf{Y}_1^{\mathrm{N}}) - I(M_1;\mathbf{Y}_2^{\mathrm{N}}) + 2\mathrm{N}\epsilon \\
&= \; \sum_{n=1}^{\mathrm{N}} [I(M_1;Y_{1,n}|\mathbf{Y}_{1,n+1}^{\mathrm{N}}) - I(M_1;Y_{2,n}|\mathbf{Y}_2^{n-1})] + 2\mathrm{N}\epsilon \\
&\overset{(c)}{=} \; \sum_{n=1}^{\mathrm{N}} [I(M_1,\mathbf{Y}_2^{n-1};Y_{1,n}|\mathbf{Y}_{1,n+1}^{\mathrm{N}}) - I(M_1,\mathbf{Y}_{1,n+1}^{\mathrm{N}};Y_{2,n}|\mathbf{Y}_2^{n-1})] + 2\mathrm{N}\epsilon \\
&\overset{(d)}{=} \; \sum_{n=1}^{\mathrm{N}} [I(M_1;Y_{1,n}|\mathbf{Y}_{1,n+1}^{\mathrm{N}},\mathbf{Y}_2^{n-1}) - I(M_1;Y_{2,n}|\mathbf{Y}_{1,n+1}^{\mathrm{N}},\mathbf{Y}_2^{n-1})] + 2\mathrm{N}\epsilon \\
&\leq \; \sum_{n=1}^{\mathrm{N}} [I(M_1,W_n;Y_{1,n}|\mathbf{Y}_{1,n+1}^{\mathrm{N}},\mathbf{Y}_2^{n-1}) - I(M_1;Y_{2,n}|\mathbf{Y}_{1,n+1}^{\mathrm{N}},\mathbf{Y}_2^{n-1})] + 2\mathrm{N}\epsilon \\
&\overset{(e)}{=} \; \sum_{n=1}^{\mathrm{N}} [I(M_1;Y_{1,n}|\mathbf{Y}_{1,n+1}^{\mathrm{N}},\mathbf{Y}_2^{n-1}) + I(W_n;Y_{1,n}|M_1,\mathbf{Y}_{1,n+1}^{\mathrm{N}},\mathbf{Y}_2^{n-1}) \\
&\qquad - I(M_1;Y_{2,n}|\mathbf{Y}_{1,n+1}^{\mathrm{N}},\mathbf{Y}_2^{n-1})] + 2\mathrm{N}\epsilon \\
&= \; \sum_{n=1}^{\mathrm{N}} [I(M_1;Y_{1,n}|\mathbf{Y}_{1,n+1}^{\mathrm{N}},\mathbf{Y}_2^{n-1}) + H(W_n|M_1,\mathbf{Y}_{1,n+1}^{\mathrm{N}},\mathbf{Y}_2^{n-1}) \\
&\qquad - H(W_n|M_1,Y_{1,n},\mathbf{Y}_{1,n+1}^{\mathrm{N}},\mathbf{Y}_2^{n-1}) - I(M_1;Y_{2,n}|\mathbf{Y}_{1,n+1}^{\mathrm{N}},\mathbf{Y}_2^{n-1})] + 2\mathrm{N}\epsilon \\
&\leq \; \sum_{n=1}^{\mathrm{N}} [I(M_1;Y_{1,n}|\mathbf{Y}_{1,n+1}^{\mathrm{N}},\mathbf{Y}_2^{n-1}) + H(W_n|M_1,\mathbf{Y}_{1,n+1}^{\mathrm{N}},\mathbf{Y}_2^{n-1}) \\
&\qquad - I(M_1;Y_{2,n}|\mathbf{Y}_{1,n+1}^{\mathrm{N}},\mathbf{Y}_2^{n-1})] + 2\mathrm{N}\epsilon,
\end{aligned}
$$

where (a) is from Fano's inequality, (b) is from confidentiality constraints, (c) and (d) follow from Csiszár's sum identity and (e) is the chain rule for mutual information. Letting $U_n \triangleq (\mathbf{Y}_{1,n+1}^{\mathrm{N}}, \mathbf{Y}_2^{n-1})$; and $V_{1,1} = \cdots = V_{1,\mathrm{N}} \triangleq M_1$, where U and V_1 satisfy the Markov chain $U \to V_1 \to X$ specified in Section 3 for the channel C_3, we get

$$
\mathrm{N}R_1 \;\leq\; \sum_{n=1}^{\mathrm{N}} [I(V_{1,n};Y_{1,n}|U_n) + H(W_n|U_n,V_{1,n}) - I(V_{1,n};Y_{2,n}|U_n)] + 2\mathrm{N}\epsilon. \tag{53}
$$

Proceeding in a similar fashion and letting $V_{2,1} = \cdots = V_{2,\mathrm{N}} \triangleq M_2$,

$$
\mathrm{N}R_2 \;\leq\; \sum_{n=1}^{\mathrm{N}} [I(V_{2,n};Y_{2,n}|U_n) + H(W_n|U_n,V_{2,n}) - I(V_{2,n};Y_{1,n}|U_n)] + 2\mathrm{N}\epsilon. \tag{54}
$$

For the channel C_3, we also derive a genie-aided outer bound by letting a hypothetical genie give D_1 message M_2, while D_2 computes the equivocation using M_2 as side-information. $\forall\epsilon > 0$ and sufficiently small; and for large N, R_1 can be upper bounded as follows:

$$
\begin{aligned}
\mathrm{N}R_1 &= H(M_1) \leq H(M_1|\mathbf{Y}_2^{\mathrm{N}}) + \mathrm{N}\epsilon \leq H(M_1, M_2|\mathbf{Y}_2^{\mathrm{N}}) + \mathrm{N}\epsilon \\
&= H(M_1|\mathbf{Y}_2^{\mathrm{N}}, M_2) + H(M_2|\mathbf{Y}_2^{\mathrm{N}}) + \mathrm{N}\epsilon \leq H(M_1|\mathbf{Y}_2^{\mathrm{N}}, M_2) + \mathrm{N}\epsilon \\
&\leq H(M_1|\mathbf{Y}_2^{\mathrm{N}}, M_2) - H(M_1|\mathbf{Y}_1^{\mathrm{N}}) + \mathrm{N}\epsilon \overset{(a)}{\leq} H(M_1|\mathbf{Y}_2^{\mathrm{N}}, M_2) - H(M_1|\mathbf{Y}_1^{\mathrm{N}}, M_2) + \mathrm{N}\epsilon \\
&\leq I(M_1; \mathbf{Y}_1^{\mathrm{N}}|M_2) - I(M_1; \mathbf{Y}_2^{\mathrm{N}}|M_2) + 2\mathrm{N}\epsilon \\
&= \sum_{n=1}^{\mathrm{N}} [I(M_1; Y_{1,n}|\mathbf{Y}_{1,n+1}^{\mathrm{N}}, M_2) - I(M_1; Y_{2,n}|\mathbf{Y}_2^{n-1}, M_2)] + 2\mathrm{N}\epsilon \\
&\overset{(b)}{=} \sum_{n=1}^{\mathrm{N}} [I(M_1, \mathbf{Y}_2^{n-1}; Y_{1,n}|\mathbf{Y}_{1,n+1}^{\mathrm{N}}, M_2) - I(M_1, \mathbf{Y}_{1,n+1}^{\mathrm{N}}; Y_{2,n}|\mathbf{Y}_2^{n-1}, M_2)] + 2\mathrm{N}\epsilon \\
&\overset{(c)}{=} \sum_{n=1}^{\mathrm{N}} [I(M_1; Y_{1,n}|\mathbf{Y}_{1,n+1}^{\mathrm{N}}, \mathbf{Y}_2^{n-1}, M_2) - I(M_1; Y_{2,n}|\mathbf{Y}_{1,n+1}^{\mathrm{N}}, \mathbf{Y}_2^{n-1}, M_2)] + 2\mathrm{N}\epsilon \\
&\leq \sum_{n=1}^{\mathrm{N}} [I(M_1, W_n; Y_{1,n}|\mathbf{Y}_{1,n+1}^{\mathrm{N}}, \mathbf{Y}_2^{n-1}, M_2) - I(M_1; Y_{2,n}|\mathbf{Y}_{1,n+1}^{\mathrm{N}}, \mathbf{Y}_2^{n-1}, M_2)] + 2\mathrm{N}\epsilon \\
&= \sum_{n=1}^{\mathrm{N}} [I(M_1; Y_{1,n}|\mathbf{Y}_{1,n+1}^{\mathrm{N}}, \mathbf{Y}_2^{n-1}, M_2) + I(W_n; Y_{1,n}|M_1, \mathbf{Y}_{1,n+1}^{\mathrm{N}}, \mathbf{Y}_2^{n-1}, M_2) \\
&\quad - I(M_1; Y_{2,n}|\mathbf{Y}_{1,n+1}^{\mathrm{N}}, \mathbf{Y}_2^{n-1}, M_2)] + 2\mathrm{N}\epsilon \\
&= \sum_{n=1}^{\mathrm{N}} [I(M_1; Y_{1,n}|\mathbf{Y}_{1,n+1}^{\mathrm{N}}, \mathbf{Y}_2^{n-1}, M_2) + H(W_n|M_1, Y_{n+1}^{\mathrm{N}}, \mathbf{Y}_2^{n-1}, M_2) \\
&\quad - H(W_n|M_1, Y_{1,n}, \mathbf{Y}_{1,n+1}^{\mathrm{N}}, \mathbf{Y}_2^{n-1}, M_2) - I(M_1; Y_{2,n}|\mathbf{Y}_{1,n+1}^{\mathrm{N}}, \mathbf{Y}_2^{n-1}, M_2)] + 2\mathrm{N}\epsilon \\
&\leq \sum_{n=1}^{\mathrm{N}} [I(M_1; Y_{1,n}|\mathbf{Y}_{1,n+1}^{\mathrm{N}}, \mathbf{Y}_2^{n-1}, M_2) + H(W_n|M_1, \mathbf{Y}_{1,n+1}^{\mathrm{N}}, \mathbf{Y}_2^{n-1}, M_2) \\
&\quad - I(M_1; Y_{2,n}|\mathbf{Y}_{1,n+1}^{\mathrm{N}}, \mathbf{Y}_2^{n-1}, M_2)] + 2\mathrm{N}\epsilon,
\end{aligned}
$$

where (a) follows since the genie gives D_1 message M_2, (b) and (c) follow from Csiszár's sum identity. Letting $U_n \triangleq (\mathbf{Y}_{1,n+1}^{\mathrm{N}}, \mathbf{Y}_2^{n-1})$, $V_{1,1} = \cdots = V_{1,\mathrm{N}} \triangleq M_1$ and $V_{2,1} = \cdots = V_{2,\mathrm{N}} \triangleq M_2$, where U, V_1 and V_2 satisfy the Markov chains $U \to V_1 \to X$ and $U \to V_2 \to X$ specified in Section 3 for the channel C_3, R_1 can be bounded as

$$\mathrm{N}R_1 \leq \sum_{n=1}^{\mathrm{N}} [I(V_{1,n}; Y_{1,n}|U_n, V_{2,n}) + H(W_n|U_n, V_{1,n}, V_{2,n}) - I(V_{1,n}; Y_{2,n}|U_n, V_{2,n})] + 2\mathrm{N}\epsilon. \quad (55)$$

Similarly,

$$\mathrm{N}R_1 \leq \sum_{n=1}^{\mathrm{N}} [I(V_{2,n}; Y_{2,n}|U_n, V_{1,n}) + H(W_n|U_n, V_{1,n}, V_{2,n}) - I(V_{2,n}; Y_{1,n}|U_n, V_{1,n})] + 2\mathrm{N}\epsilon. \quad (56)$$

For the channel C$_3$, the outer bound on R_1+R_2 can be made tighter by the following procedure. From (19) - (20), we see that

$$R_1+R_2 \le I_1+I_2, \tag{57}$$

$$R_1+R_2 \le I_1^*+I_2^*. \tag{58}$$

Therefore,

$$R_1+R_2 \le \min[I_1+I_2^*, I_2+I_1^*]. \tag{59}$$

We show now that the bound (59) is a tighter bound than (57) and (58). It is easy to see that

$$I_1+I_2 = I_1^*+I_2^*+I(W;V_1|U,V_2)+I(W;V_2|U,V_1).$$

Consider $2(I_1+I_2) = 2[I_1^*+I_2^*+I(W;V_1|U,V_2)+I(W;V_2|U,V_1)]$, which implies the following:

$$\begin{aligned}
\min[I_1+I_2^*, I_2+I_1^*] &\le I_1+I_2,\\
\min[I_1+I_2^*, I_2+I_1^*] &\le I_1^*+I_2^*.
\end{aligned}$$

Therefore, the sum rate bound given by (59) is tighter than (57) and (58).

5.5 Proof of Theorem 3.8

For the channel C$_3$, $\forall\epsilon>0$ and sufficiently small; and for large N, R_1 can be bounded as follows:

$$\begin{aligned}
\mathrm{N}R_1 &= H(M_1) = I(M_1;\mathbf{Y}_1^{\mathrm{N}}) + H(M_1|\mathbf{Y}_1^{\mathrm{N}})\\
&\overset{(a)}{\le} I(M_1;\mathbf{Y}_1^{\mathrm{N}}) + \mathrm{N}\epsilon \overset{(b)}{\le} I(M_1;\mathbf{Y}_1^{\mathrm{N}}) - I(M_1;\mathbf{Y}_2^{\mathrm{N}}) + 2\mathrm{N}\epsilon,
\end{aligned}$$

where (a) follows from Fano's inequality; and (b) follows from confidentiality constraints. Following the procedure used to prove Theorem 3.3 (see Section 5.2) and letting $U_{\mathrm{n}} = (\mathbf{W}_{\mathrm{n}+1}^{\mathrm{N}}, \mathbf{Y}_1^{\mathrm{n}-1}, \mathbf{Y}_{2,\mathrm{n}+1}^{\mathrm{N}})$ and $V_{1,\mathrm{n}} = M_1$,

$$\mathrm{N}R_1 \le \sum_{\mathrm{n}=1}^{\mathrm{N}} I(U_{\mathrm{n}}, V_{1,\mathrm{n}}; Y_{1,\mathrm{n}}) - I(V_{1,\mathrm{n}}; W_{\mathrm{n}}|U_{\mathrm{n}}) - I(V_{1,\mathrm{n}}; Y_{2,\mathrm{n}}) + 2\mathrm{N}\epsilon. \tag{60}$$

Similarly, letting $V_{2,\mathrm{n}} = M_2$, we get

$$\mathrm{N}R_2 \le \sum_{\mathrm{n}=1}^{\mathrm{N}} I(U_{\mathrm{n}}, V_{2,\mathrm{n}}; Y_{2,\mathrm{n}}) - I(V_{2,\mathrm{n}}; W_{\mathrm{n}}|U_{\mathrm{n}}) - I(V_{2,\mathrm{n}}; Y_{1,\mathrm{n}}) + 2\mathrm{N}\epsilon, \tag{61}$$

and the following bounds on the sum-rate $R_1 + R_2$:

$$\mathrm{N}(R_1+R_2) \leq \sum_{n=1}^{\mathrm{N}}[I(U_{\mathrm{n}}, V_{1,\mathrm{n}}; Y_{1,\mathrm{n}}) - I(V_{1,\mathrm{n}}; W_{\mathrm{n}}|U_{\mathrm{n}})]$$
$$+\sum_{\mathrm{n}=1}^{\mathrm{N}}[I(U_{\mathrm{n}}, V_{2,\mathrm{n}}; Y_{2,\mathrm{n}}|V_{1,\mathrm{n}}) - I(V_{2,\mathrm{n}}; W_{\mathrm{n}}|V_{1,\mathrm{n}}, U_{\mathrm{n}})] - I(V_{1,\mathrm{n}}; Y_{2,\mathrm{n}}) + 2\mathrm{N}\epsilon, \tag{62}$$
$$\mathrm{N}(R_1+R_2) \leq \sum_{\mathrm{n}=1}^{\mathrm{N}}[I(U_{\mathrm{n}}, V_{2,\mathrm{n}}; Y_{2,\mathrm{n}}) - I(V_{2,\mathrm{n}}; W_{\mathrm{n}}|U_{\mathrm{n}})]$$
$$+\sum_{\mathrm{n}=1}^{\mathrm{N}}[I(U_{\mathrm{n}}, V_{1,\mathrm{n}}; Y_{1,\mathrm{n}}|V_{2,\mathrm{n}}) - I(V_{1,\mathrm{n}}; W_{\mathrm{n}}|V_{2,\mathrm{n}}, U_{\mathrm{n}})] - I(V_{2,\mathrm{n}}; Y_{1,\mathrm{n}}) + 2\mathrm{N}\epsilon. \tag{63}$$

A time sharing RV Q, which is uniformly distributed over N symbols and independent of the RVs M_1, M_2, W, U, V_1, V_2, X, Y_1 and Y_2 is introduced for the single letter characterization of the above derived outer bounds. Applying the procedure similar to the one presented in [24, Chapter 15.3.4] on the N-letter expressions obtained in the above stated theorems, we get the outer bounds presented in Section 3. This completes the proofs of Theorem 3.2, Theorem 3.3, Theorem 3.5, Theorem 3.7 and Theorem 3.8.

6 Conclusions

We presented inner and outer bounds on the capacity region of three classes of two-user discrete memoryless broadcast channels, with noncausal side-information at the encoder. We generalized existing approaches to prove the achievability theorems, and characterized the rate penalties for having to deal with side-information at the encoder. For channels with confidentiality constraints, we showed that rate penalties exist for dealing with both side-information and confidentiality constraints. In the case of outer bounds, we focus on the explicit characterization of the sum-rate bounds. For channels where each receiver has *a priori* knowledge of the message of the other receiver, we showed that the outer bounds are only a factor away from the achievable region, where the factor is independent of the channel distribution.

A Encoder error analysis

Here, we upper bound the probability of encoder error for the channel C_1, by using results from the second moment method [22]. This method was also employed in [26] and [27, Chap. 7, pp. 354] to provide an alternative proof of Marton's achievability scheme. An error is declared at the encoder of S if it is not possible to find a pair (i_1, i_2) to satisfy the condition $E_{\mathrm{S}} \triangleq \{(\mathbf{W}, \mathbf{V}_1(i_1, j_1), \mathbf{V}_2(i_2, j_2)) \in A_\epsilon^{(\mathrm{N})}(P_{W,V_1,V_2})\}$. Let $P_{e,E_{\mathrm{S}}}$ denote the probability of error at the encoder, *i.e.*, $P_{e,E_{\mathrm{S}}} \triangleq \Pr(E_{\mathrm{S}}^c)$. Let I be an indicator RV that the event E_{S} has occurred. Let $Q = \sum_{j_1,j_2} I$; $\bar{Q} = \mathbb{E}[Q]$; and $\mathrm{Var}[Q] = \mathbb{E}[(Q - \bar{Q})^2]$, where $\mathbb{E}(.)$ denotes the expectation operator. $P_{e,E_{\mathrm{S}}}$ can be upper bounded as follows:

$$P_{e,E_{\mathrm{S}}} = \Pr(Q = 0) \overset{(i)}{\leq} \mathrm{Var}[Q]/\bar{Q}^2, \tag{64}$$

where (i) follows from Markov's inequality for non-negative RVs. Consider now

$$\begin{aligned}\bar{Q} &= \sum_{j_1,j_2} \mathbb{E}(I) \geq \sum_{j_1,j_2} (1 - \delta^{(\mathrm{N})}) 2^{-\mathrm{N}[I(V_1;V_2)+I(V_1,V_2;W)+4\epsilon]} \\ &= (1 - \delta^{(\mathrm{N})}) 2^{-\mathrm{N}[R_1^* + R_2^* - I(V_1;V_2) - I(V_1,V_2;W) - 4\epsilon]}.\end{aligned}$$

Next, consider $\mathrm{Var}[Q] = \sum_{j_1,j_2} \sum_{j_1',j_2'} \{\mathbb{E}[I(j_1,j_2)I(j_1',j_2')] - \mathbb{E}[I(j_1,j_2)]\mathbb{E}I(j_1',j_2')]\}$. We have the following four cases:

1. If $j_1' \neq j_1$ and $j_2' \neq j_2$, then $I(j_1,j_2)$ and $I(j_1',j_2')$ are independent and $\mathrm{Var}[Q] = 0$.

2. If $j_1' = j_1$ and $j_2' = j_2$, then $\mathbb{E}[I(j_1,j_2)I(j_1',j_2')] = \mathbb{E}[I(j_1,j_2)] \leq 2^{-\mathrm{N}[I(V_1;V_2)+I(V_1,V_2;W)-4\epsilon]}$.

3. If $j_1' \neq j_1$ and $j_2' = j_2$, then $\mathbb{E}[I(j_1,j_2)I(j_1',j_2')] \leq 2^{-\mathrm{N}[I(V_1;V_2|U)+I(V_1,V_2;W)+I(V_1;V_2,W)-6\epsilon]}$.

4. If $j_1' = j_1$ and $j_2' \neq j_2$, then $\mathbb{E}[I(j_1,j_2)I(j_1',j_2')] \leq 2^{-\mathrm{N}[I(V_1;V_2|U)+I(V_1,V_2;W)+I(V_2;V_1,W)-6\epsilon]}$.

Substituting for $\bar{Q}$ and $\mathrm{Var}[Q]$ in (64), we can show that $P(E_{\mathrm{S}}) \leq \delta_{\mathrm{C}_1}^{(\mathrm{N})}$, $\forall \delta_{\mathrm{C}_1}^{(\mathrm{N})} > 0$ and sufficiently small; and for N large, if the following conditions are simultaneously satisfied:

$$\begin{aligned} R_1^{'} &> I(W;V_1) - \epsilon_1, & (65)\\ R_2^{'} &> I(W;V_2) - \epsilon_2, & (66)\\ R_1^{'} + R_2^{'} &> I(V_1;V_2) + I(V_1,V_2;W) - \epsilon_3. & (67)\end{aligned}$$

Similar analysis results in a bound on the binning rates for the channel C_3. The probability of encoder error $P(E_{\mathrm{S}}) \leq \delta_{\mathrm{C}_3}^{(\mathrm{N})}$, $\forall \delta_{\mathrm{C}_3}^{(\mathrm{N})} > 0$ and sufficiently small; and for N large, if the following conditions are simultaneously satisfied:

$$\begin{aligned} R_1^* &> I(W;V_1|U)-\epsilon_{11}, \quad (68)\\ R_2^* &> I(W;V_2|U)-\epsilon_{12}, \quad (69)\\ R_1^*+R_2^* &> I(V_1;V_2|U)+I(V_1,V_2;W|U)-\epsilon_{13}. \quad (70)\end{aligned}$$

References

[1] T. Cover, "Broadcast channels," *IEEE Trans. Inf. Theory*, vol. 18, no. 1, pp. 2–14, Jan. 1972.

[2] K. Marton, "A coding theorem for the discrete memoryless broadcast channel," *IEEE Trans. Inf. Theory*, vol. 25, no. 3, pp. 306–311, May 1979.

[3] A. Gohari, A. El Gamal, and V. Anantharam, "On an outer bound and an inner bound for the general broadcast channel," in *Proc. IEEE Int. Symp. Inf. Theory*, Jun. 2010, pp. 540–544.

[4] H. Sato, "An outer bound to the capacity region of broadcast channels (Corresp.)," *IEEE Trans. Inf. Theory*, vol. 24, no. 3, pp. 374–377, May 1978.

[5] C. Nair and A. El Gamal, "An outer bound to the capacity region of the broadcast channel," *IEEE Trans. Inf. Theory*, vol. 53, no. 1, pp. 350–355, Jan. 2007.

[6] A. E. Gamal, "The capacity of a class of broadcast channels," *IEEE Trans. Inf. Theory*, vol. 25, no. 2, pp. 166–169, Mar. 1979.

[7] C. Nair, "A note on outer bounds for broadcast channel," in *Proc. Int. Zurich Seminar Comm.*, 2010. [Online]. Available: http://arxiv.org/abs/1101.0640v1

[8] Y. Steinberg and S. Shamai (Shitz), "Achievable rates for the broadcast channel with states known at the transmitter," in *Proc. IEEE Int. Symp. Inf. Theory*, Adelaide, SA, Sep. 2005, pp. 2184–2188.

[9] Y. Steinberg, "Coding for the degraded broadcast channel with random parameters, with causal and noncausal side information," *IEEE Trans. Inf. Theory*, vol. 51, no. 8, pp. 2867–2877, Aug. 2005.

[10] G. Kramer and S. Shamai (Shitz), “Capacity for classes of broadcast channels with receiver side information,” in *Proc. IEEE Inf. Theory Workshop*, Tahoe City, CA, Sep. 2007, pp. 313–318.

[11] Y. Liang, H. V. Poor, and S. Shamai (Shitz), “Information theoretic security,” *Found. Trends Comm. Inf. Theory*, vol. 5, no. 4, pp. 355–580, Apr. 2009.

[12] I. Csiszár and J. Körner, “Broadcast channels with confidential messages,” *IEEE Trans. Inf. Theory*, vol. IT-24, no. 3, pp. 339–348, May 1978.

[13] E. Ekrem and S. Ulukuş, “Secrecy capacity of a class of broadcast channels with an eavesdropper,” *EURASIP J. Wireless Comm. and Net.*, vol. 2009, Article ID 824235, 29 pages, 2009, doi: 10.1155/2009/824235.

[14] R. Liu, I. Marić, P. Spasojević, and R. D.Yates, “Discrete memoryless interference and broadcast channels with confidential messages: Secrecy rate regions,” *IEEE Trans. Inf. Theory*, vol. 54, no. 6, pp. 2493–2507, Jun. 2008.

[15] G. Bagherikaram, A. Motahari, and A. Khandani, “Secrecy capacity region of Gaussian broadcast channel,” in *Proc. IEEE* 43^{rd} *Annual Conf. Inf. Sciences Syst.*, Baltimore, MD, Mar. 2009, pp. 152–157.

[16] R. Liu and H. V. Poor, “Secrecy capacity region of a multiple-antenna Gaussian broadcast channel with confidential messages,” *IEEE Trans. Inf. Theory*, vol. 55, no. 3, pp. 1235–1249, Mar. 2009.

[17] H. Ly, T. Liu, and Y. Liang, “Multiple-input multiple-output Gaussian broadcast channels with common and confidential messages,” *IEEE Trans. Inf. Theory*, vol. 56, no. 11, pp. 5477–5487, Nov. 2010.

[18] Y. Liang, H. V. Poor, and S. Shamai (Shitz), “Physical layer security in broadcast networks,” *Security Comm. Net.*, vol. 2, no. 3, pp. 227–238, May/Jun. 2009.

[19] U. Maurer and S. Wolf, “Information-theoretic key agreement: From weak to strong secrecy for free,” in *Proc.* 19^{th} *Int. Conf. Theory App. Crypt. Tech.*, Bruges, Belgium, 2000, pp. 351–368.

[20] S. Gel’fand and M. Pinsker, “Coding for channels with random parameters,” *Probl. Contr. and Inf. Theory*, vol. 9, no. 1, pp. 19–31, 1980.

[21] Y. Chen and A. J. Han Vinck, "Wiretap channel with side information," *IEEE Trans. Inf. Theory*, vol. 54, no. 1, pp. 395–402, Jan. 2008.

[22] N. Alon and J. H. Spencer, *The Probabilistic Method*, 2nd ed. New York: John Wiley, 2000.

[23] T. Oechtering and M. Skoglund, "Coding for the bidirectional broadcast channel with random states known at the encoder," in *Proc. IEEE Int. Symp. Inf. Theory*, Seoul, S.Korea, Jul. 2009, pp. 2013–2017.

[24] T. Cover and J. Thomas, *Elements of Information Theory*, 2nd ed. New York: Wiley-Interscience, 2006.

[25] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems.* Orlando, FL, USA: Academic Press, Inc., 1982.

[26] A. E. Gamal and E. C. van der Meulen, "A proof of Marton's coding theorem for the discrete memoryless broadcast channel," *IEEE Trans. Inf. Theory*, vol. 27, no. 1, pp. 120–122, Jan. 1981.

[27] G. Kramer, "Topics in multi-user information theory," *Found. Trends Comm. Inf. Theory*, vol. 4, no. 4-5, pp. 265–444, 2007. [Online]. Available: http://ee.usc.edu/~gkramer/Papers/kramerNOW07.pdf

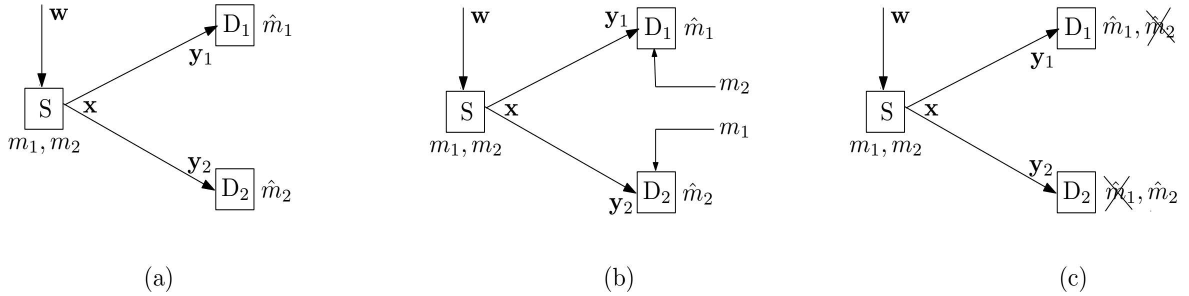

Figure 1: State-dependent broadcast channels with side-information at the transmitter: (a) Class I; (b) Class II; and (c) Class III.

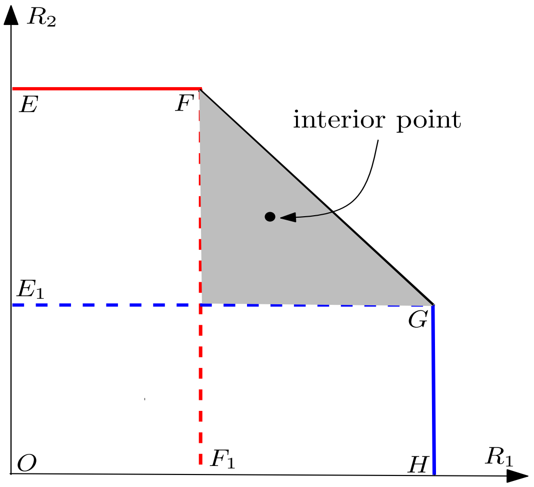

Figure 2: Pictorial representation of the rate region for Class I channels.

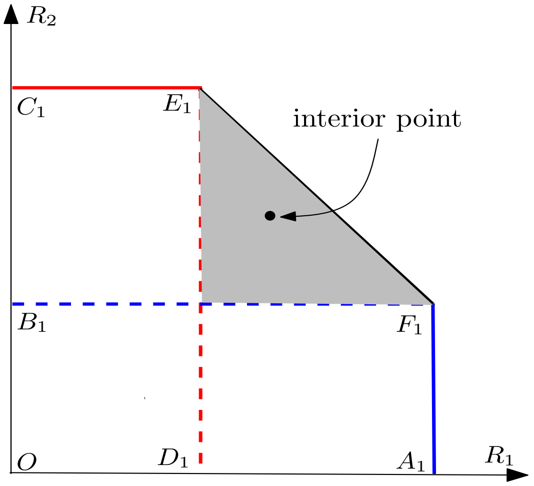

Figure 3: Pictorial representation of the rate region for Class III channels.